

Audiovisual associations alter the perception of low-level visual motion

Hulusi Kafaligonul^{1*} and Can Oluk²

¹ National Magnetic Resonance Research Center (UMRAM), Bilkent University, Ankara, Turkey, ² Department of Psychology, Bilkent University, Ankara, Turkey

Motion perception is a pervasive nature of vision and is affected by both immediate pattern of sensory inputs and prior experiences acquired through associations. Recently, several studies reported that an association can be established quickly between directions of visual motion and static sounds of distinct frequencies. After the association is formed, sounds are able to change the perceived direction of visual motion. To determine whether such rapidly acquired audiovisual associations and their subsequent influences on visual motion perception are dependent on the involvement of higher-order attentive tracking mechanisms, we designed psychophysical experiments using regular and reverse-phi random dot motions isolating low-level pre-attentive motion processing. Our results show that an association between the directions of low-level visual motion and static sounds can be formed and this audiovisual association alters the subsequent perception of low-level visual motion. These findings support the view that audiovisual associations are not restricted to high-level attention based motion system and early-level visual motion processing has some potential role.

Keywords: motion perception, visual motion processing, audiovisual associations, multisensory, direction discrimination

OPEN ACCESS

Edited by:

Sharon Gilaie-Dotan,
University College London, UK

Reviewed by:

Souta Hidaka,
Rikkyo University, Japan
Akitoshi Ogawa,
University of Tokyo, Japan

*Correspondence:

Hulusi Kafaligonul,
National Magnetic Resonance
Research Center (UMRAM), Bilkent
University, Bilkent Cyberpark,
Cyberplaza, Ankara 06800, Turkey
hulusi@bilkent.edu.tr

Received: 28 October 2014

Accepted: 14 March 2015

Published: 31 March 2015

Citation:

Kafaligonul H and Oluk C (2015)
Audiovisual associations alter the
perception of low-level visual motion.
Front. Integr. Neurosci. 9:26.
doi: 10.3389/fnint.2015.00026

Introduction

Perception is shaped by both immediate pattern of sensory inputs and our prior experiences with the external world. Phenomena such as adaptation and learning have been used extensively to understand how prior experiences over different timescales are involved in shaping visual motion perception. Of particular interest is that associative learning (the process by which an association between two stimuli is learned) revealed unexpected levels of sensory plasticity and changes in visual motion processing (Schlack and Albright, 2007; Schlack et al., 2007). However, these studies mostly focused on a single sensory modality (i.e., vision). As learning in natural settings typically involves multisensory input, such unisensory learning might not be optimal for engaging learning mechanisms adapted to operate in multisensory environments. Consistent with this ecological argument, multisensory associations have, in a series of recent studies, been discovered to have dramatic influences on the perception of visual motion (for a review see Shams et al., 2011).

A recent paradigm developed by Teramoto et al. (2010) provides compelling evidence that multisensory associations influence visual motion perception¹. In their study, two circles were

¹ Later studies called this paradigm as “sound contingent motion aftereffect” (Hidaka et al., 2011; Kobayashi et al., 2012a,b). In this study, we preferred to use the term “audiovisual associations” due to the following reasons: (i) aftereffects usually result in a bias opposite to the exposed direction and they are considered as a product of the recalibration of the sensory system, (ii) as pointed out in several studies (Seitz et al., 2007; Kim et al., 2009; Albright, 2012; Murphy et al., 2014), associations can be established implicitly (i.e., without any conditional reinforcement) at short time-scales.

flashed side by side with a temporal offset at the right of a fixation target. The flashed circles induced a perception of apparent motion (i.e., a circle moving either from left-to-right or right-to-left) when the spatial and temporal offset between them were small. Expectedly, a static flickering circle was perceived if the circles were presented with a temporal offset at the same spatial location. During a typical association phase, observers viewed these apparent motion displays alternating between rightward and leftward directions. Each flashed circle of apparent motion was synchronized with a brief static sound of distinct frequency. The static sounds led to two different sound sequences and each one of these sequences was paired with a motion direction. After the association phase, the sounds biased the perceived direction of apparent motion and even induced motion perception of a static flickering circle in the direction as previously paired with the given sound sequence during the association phase. Accordingly, brief static sounds determined the perceived visual motion direction and this driving effect acquired through associations lasted for at least a few days. Using a similar paradigm, Hidaka et al. (2011) showed that, after an audiovisual association phase, static sounds are also able to change the perceived direction of random dot displays. As opposed to previous associative learning studies (e.g., Haijiang et al., 2006), these studies provide a clear demonstration that associations between static sounds and visual motion direction can be established in a short amount of time and they do not require any conditional reinforcement. Moreover, these audiovisual associations have been found to be selective for certain stimulus features (e.g., visual field, sound frequency), suggesting that they occur at the perceptual level rather than any decision level. However, the neural mechanisms and processing stages underlying them remain unclear within the well studied visual motion hierarchy. An obvious question to ask is whether such associations are restricted to later stages of visual motion processing and only mediated by higher-order sensory areas.

The existence of two basic motion systems is suggested by many studies². First one, low-level pre-attentive motion energy system, is dependent on the spatiotemporal (positional) changes in luminance (Anstis, 1980; Braddick, 1980; Adelson and Bergen, 1985). The low-level motion energy is believed to be computed by the motion detectors located at lower motion areas. Moreover, a number of studies have shown that attentive tracking of salient visual features can overcome ambiguity in low-level spatio-temporal cues and generate consistent visual motion percept (Lu and Sperling, 1995a; Culham et al., 2000). It has been proposed that attentive tracking constitutes a high-level attention based motion system mediated by higher-order motion areas (Cavanagh, 1992; Lu and Sperling, 1995b). As mentioned above, in order to create motion, previous audiovisual association studies on visual motion perception used either a single visual object or random dots not very well controlled for tracking of displaced features. Such stimuli can be localizable in space and activate both low-level motion energy and high-level attention-dependent

positional tracking mechanisms. In this respect, they do not provide any specific information about the neural mechanisms underlying audiovisual associations.

In the current study, we aimed to determine whether audiovisual associations could be formed and affect visual motion perception in absence of higher-order attentive tracking mechanisms. If so, this outcome would imply that audiovisual associations are not restricted to later stages of visual motion processing (i.e., high-level attention based motion system). We took two basic approaches to minimize the involvement of higher-order attentive tracking and to isolate low-level pre-attentive visual motion processing. The involvement of high-level processes can be changed by manipulating the parameters of random-dot displays (Sato, 1998). For instance, when the dot size is decreased and the dot density is increased, it becomes difficult to attentively track the position of individual dots and hence, high-level processes become less effective (Ho and Giaschi, 2009a,b). Therefore, we first optimized the parameters of our random dot displays to minimize any explicit attention-dependent position tracking of individual dots. In addition, using reverse-phi illusion has also been found to be a fruitful approach to restrict the involvement of higher order processes (Azzopardi and Hock, 2011; Kafaligonul and Stoner, 2012; Hock and Nichols, 2013). In this illusion, perceived direction of motion is opposite to the physical displacement and consistent with the outcome of low-level motion energy mechanisms rather than high-level attentive tracking of displaced features (Anstis, 1970). Accordingly, we also used reverse-phi random dot motions in our psychophysical experiments.

Materials and Methods

Participants

Ten observers (four females, six males; age range: 21–32 years) participated in this study with eight being naive to the purpose of the experiments. There were five observers in Experiment 1 and six observers in Experiment 2. Only one of the naive observers took part in both experiments. All participants had normal hearing and normal or corrected-to-normal visual acuity. Participants gave informed consent, and all procedures were in accordance with international standards (Declaration of Helsinki) and approved by the ethics committee at Faculty of Medicine, Ankara University.

Apparatus and Stimuli

We used Matlab version 7.12 (Mathworks) with Psychtoolbox 3.0 (Brainard, 1997; Pelli, 1997) for stimulus presentation and data acquisition. The precision of visual and auditory stimuli was achieved by high-performance video (NVIDIA GeForce Graphics) and sound (Asus Xonar ASIO Compatible) cards. Visual stimuli were presented on a 21" LCD monitor (NEC MultiSync 2190UXp, 1600 × 1200 pixel resolution and 60 Hz refresh rate) at a viewing distance of 57 cm. A SpectroCAL photometer was used for luminance calibration and gamma correction of the display. Sounds were emitted by headphones (Sennheiser HD 518) and amplitudes were measured by a

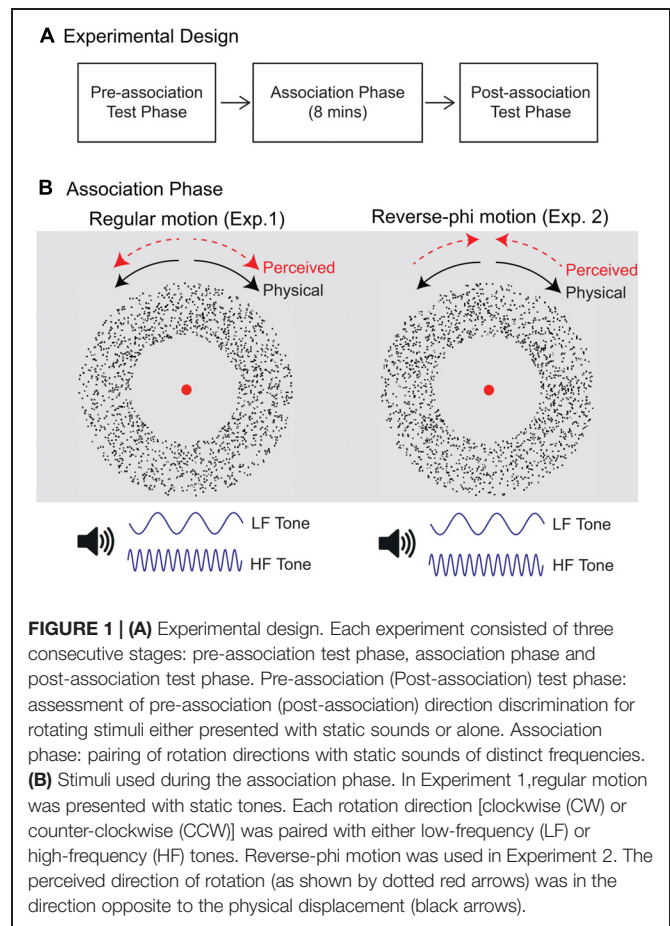
² An additional motion mechanism (second-order motion system), which is essential to extract the motion of features not defined by differences in luminance, is also supported by theoretical and experimental evidence (Lu and Sperling, 1995b).

sound-level meter (SL-4010 Lutron). Head movements were constrained by a chin rest. All experiments were performed in a dark room.

A small red circle (12 arc-min diameter) at the center of the display served as a fixation target. Visual motion stimuli were rotating dynamic random dot displays and consisted of ten random-dot frames presented within an annulus aperture (inner radius: 2.25° , outer radius: 4.5°). The dot density of each random-dot frame was 33 dots/deg and each dot had a 3.15 arc-min diameter. For regular phi motion, all dots were either lighter (51.53 cd/m^2) or darker (0.62 cd/m^2) than the gray background (26.19 cd/m^2). We used the same luminance values for reverse-phi motion. However, instead of having fixed dot luminance throughout the duration of motion, the contrast polarity of dots were reversed in each of the two consecutive random-dot frames. The duration of each random-dot frame was 100 ms and there was no temporal interval between each frame for both motion types. To control the strength of motion signal and determine the position of every dot in each random-dot frame, we used white noise motion algorithm originally developed by Britten et al. (1992). In brief, the motion signal was generated by randomly selecting a percentage of the dots ("coherent dots") from each random-dot frame to be replotted at a shifted location in a single direction [clockwise (CW) or counter-clockwise (CCW)]. The remaining dots ("noise dots") were replotted at random positions in the next random-dot frame. Even though this manipulation resulted in noise dots with random speed and direction, the coherent dots moved in the single direction with a constant speed (15° rotation per second). Based on the algorithm, the lifetime of coherent motion dots was probabilistic outcome of motion coherence (Pilly and Seitz, 2009; Schutz et al., 2010). When the proportion of coherent dots were low, it became hard to distinguish them from dynamic noise dots due to their short lifetimes. However, the lifetimes of coherent dots got longer as their proportion was increased. For the 100% coherent dots condition, all dots moved in a single direction and stayed on the display during the presentation of motion. Auditory stimulus was a static tone of either low (500 Hz) or high (2000 Hz) frequency and it was presented for 1 s (same as visual motion duration) at 83 dB SPL. Each tone burst was windowed with a 10 ms rise-and-fall time and sampled at 44.1 kHz.

Procedure

The procedure was similar to the one reported in Hidaka et al. (2011). Each experiment consisted of three separate stages: pre-association test phase, association phase, post-association test phase (Figure 1A). During each association phase, the directions of motion stimuli with 100% coherent dots were paired with static tones of distinct frequencies (Figure 1B). In Experiment 1, regular random dot motion (either lighter or darker than the background) were presented with static tones. Three observers viewed CW and CCW rotating stimuli with low (CW sound condition) and high (CCW sound condition) frequency tones, respectively. The pairing of motion directions and static tones was reversed for the remaining observers. Each direction of motion was pseudorandomly presented 120 times. Therefore, there were 240 total presentations and each



association phase (including the time interval between each presentation) lasted for around 8 min. New random dot displays were generated for each presentation. Observers were asked to attend to both rotating stimuli and static tones while they were fixating on the red circle at the center of the display.

The pre- and post-association test phase were exactly the same and each of them included two separate experimental sessions: regular and reverse-phi motion session. In each experimental session, visual motion stimuli (either regular phi or reverse-phi) rotated in either CW or CCW direction with coherence level chosen pseudorandomly from six values: 5, 15, 25, 35, 50, and 80%. Moreover, each experimental session had a balanced mixture of two auditory stimulus conditions: CW (static tone paired with CW direction during the association phase) and CCW (static tone paired with CCW direction during the association phase) sound conditions, and one visual-only (no tone) stimulus condition. Every stimulus was presented eight times per session. Accordingly, there were 288 trials ($3 \text{ audiovisual conditions} \times 6 \text{ motion coherence levels} \times 2 \text{ motion directions} \times 8 \text{ trials per stimulus}$) in each experimental session. Observers engaged in a motion direction discrimination task (two-alternative forced-choice). They fixated on the red circle at the center of the display. They were instructed neither to attend to dots (individual or groups) nor to track their displacement. They were told that

some visual motion stimuli would be accompanied by tones and asked to pay attention to both motion and tones. New random dot displays were generated for each trial to rule out tracking of previously identified dot patterns. At the end of each trial, observers indicated, by pressing left or right arrow keys, whether the moving stimuli rotated CW or CCW direction. Each observer completed four experimental sessions (2 test phases \times 2 motion types) and the order of phi and reverse-phi sessions were randomized for each testing stage. Prior to these experimental sessions, each participant was shown examples of visual-only regular and reverse-phi stimuli followed by a practice session for each motion type.

We used the same procedure and stimuli in Experiment 2 except for the motion stimuli used during the association phase. In the association phase of Experiment 2, reverse-phi motion stimuli with 100% coherent dots were used. The perceived direction (opposite to the physical displacement) of reverse-phi motion and static tones of distinct frequencies were paired (**Figure 1B**). Therefore, two sound conditions were defined based on the perceived direction of the reverse-phi motion.

Data Analysis

For auditory and visual-only conditions, a cumulative Gaussian function was fitted to the single and group-averaged data by using psignifit (version 2.5.6), a software package that implements the maximum likelihood method described by Wichmann and Hill (2001a,b). The 50% point on the resultant curves yields the point of subjective equality (PSE). PSE is the motion coherence level for which CW and CCW motion directions are equiprobable. We used the standard deviation (SD) of the fitted cumulative Gaussian as a measure of the slope of psychometric function and direction discrimination sensitivity. As the SD decreases, the psychometric curves become steeper indicating a better direction discrimination and a lower direction discrimination threshold. To quantify the effect of each static sound on PSE values, we calculated the PSE shifts by subtracting the PSE values for the visual-only condition from the ones for each sound condition. Moreover, in order to assess the sound specific changes in direction discrimination, we computed slope changes by subtracting the SD values for the visual-only condition from the ones for each sound condition. We applied a repeated measures ANOVA and follow-up tests for simple main effects on these difference values computed from individual observers.

Results

Experiment 1: Regular Motion Direction and Static Tone Association

Observers were not able to discriminate the direction of regular random dot motion when the proportion of coherent dots was low. As the proportion of coherent dots (i.e., motion coherence level) was increased, they perceived the direction of motion in the direction of physical displacement, and the direction discrimination performance of each observer was improved (**Figures 2A,B**). Motion coherence level also had a similar effect on reverse-phi

direction discrimination. However, unlike in regular motion, observers perceived the direction of reverse-phi motion in the direction opposite of the displaced dots, thus inconsistent with the tracking of dot positions (**Figures 2C,D**). The perceived direction of reverse-phi motion is instead consistent with the outcome of low-level motion energy mechanisms (Adelson and Bergen, 1985), which have been considered to be implemented within early levels of motion processing (Emerson et al., 1987; Krekelberg and Albright, 2005). Even though it is very difficult to track the position of random dots in our regular motion stimuli, the perceived direction in reverse-phi motion illusion provided us with a more explicit approach to check whether audiovisual associations affect the activity of low-level motion mechanisms.

For both motion types, static tones did not cause any change in the psychometric curves at the pre-association phase. After the paired presentation of motion directions and static tones during the association phase, the static tones significantly biased the perceived direction of regular motion in favor of the exposed audiovisual pairing. In the post-association phase, the static tone paired with the CW direction (CW sound condition) led to an increase in the CW direction reports relative to the one paired with the CCW direction (CCW sound condition). Therefore, the psychometric curve for the CW sound condition was shifted to negative coherency values and its PSE value was smaller than the CCW sound condition (**Figure 2B**). More importantly, each static tone also biased the perceived direction (opposite to the physical displacement) of reverse-phi motion in the associated direction. As in the regular motion, the CW sound condition increased the CW direction reports of the reverse-phi motion direction and the CW direction reports decreased for the CCW sound condition. This led to an increase in the PSE value of the psychometric curve relative to the one for the CCW sound condition (**Figure 2D**).

In support of the aforementioned observations above, a statistical analysis of the shifts in PSE values for each sound condition relative to the visual-only showed only a significant three-way interaction [three-way repeated-measures ANOVA motion type, test-phase and sound condition as factors, $F(1,4) = 11.132$, $p < 0.05$]. The main effects of motion type [$F(1,4) = 0.008$, $p = 0.934$], test phase [$F(1,4) = 0.547$, $p = 0.501$] and sound condition [$F(1,4) = 1.107$, $p = 0.352$] were not significant. We tested for simple main effects to understand the exact nature of the three-way interaction. For regular phi motion, the follow-up tests revealed a significant effect of sound on PSE shifts in the post-association test [**Figure 3A**, $F(1,4) = 19.655$, $p < 0.05$] but not a significant effect in the pre-association test phase [$F(1,4) = 0.834$, $p = 0.413$]. In the post-association test of reverse-phi motion, the PSE shift for CW sound condition was also significantly greater than the one for CCW sound in the expected direction [**Figure 3B**, $F(1,4) = 6.371$, $p < 0.05$]. There was no significant difference between sound conditions in the pre-association test phase [$F(1,4) = 2.362$, $p = 0.199$]. In addition, we also estimated PSE shifts for bright and dark conditions of regular motion separately. Our tests revealed non-significant effects of dot polarity and its interactions with the other factors (Supplementary Table S1).

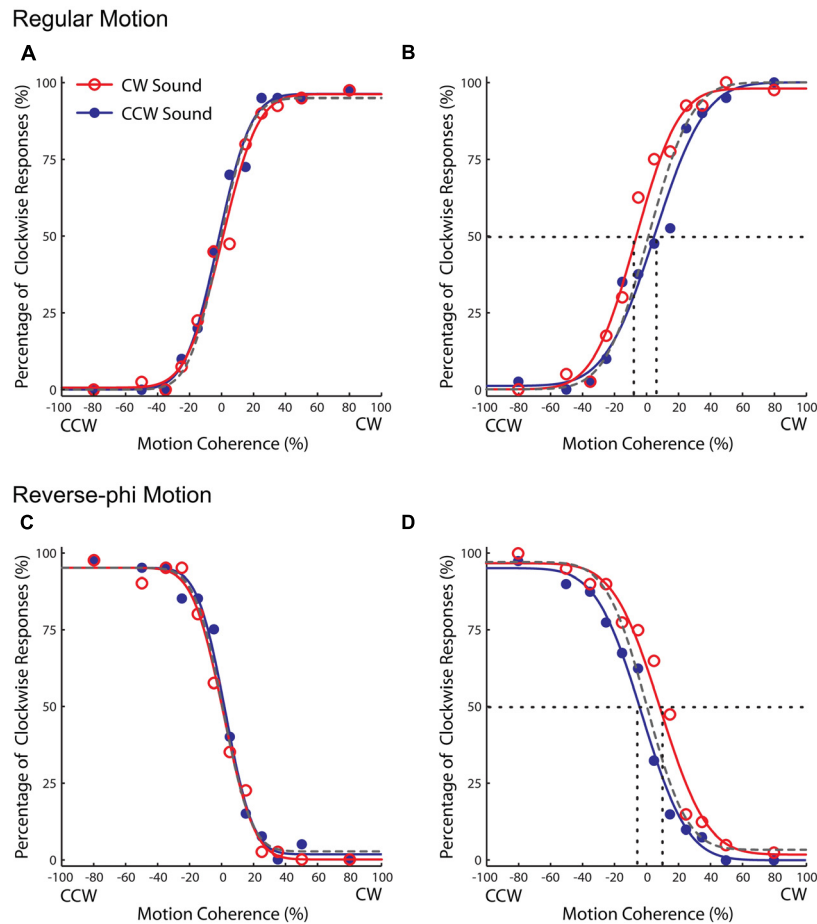


FIGURE 2 | Results of Experiment 1 (regular motion used in the association phase). Group averaged data ($n = 5$, four were naïve observers) for different experimental conditions. Top panel: pre-association (A) and post-association (B) data for regular motion. Bottom panel: pre-association (C) and post-association (D) data for reverse-phi motion. Each plot indicates the proportion of trials in which the visual stimulus was judged to rotate in

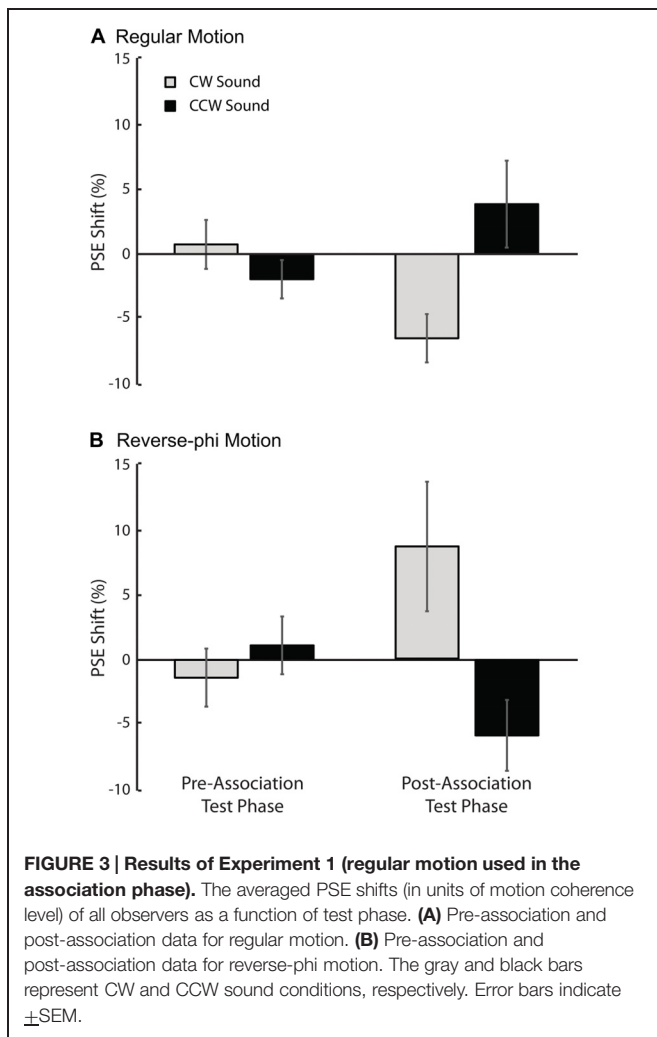
clockwise direction. The negative and positive values on the horizontal axis correspond to CCW and CW directions, respectively. The open and filled symbols represent different sound conditions. To avoid clutter, the data for the visual-only condition were not shown and the dashed gray curve corresponds to psychometric fit for that condition. The intersection of the 50% point with the vertical line gave an estimate of the PSE for each condition.

Experiment 2: Reverse-phi Motion Direction and Static Tone Association

We used 100% coherent rotating dots for the association phase of Experiment 1. Basically, the dots stayed on the display during the presentation of each motion presentation. Due to this long lifetime of dots, observers could have conceivably attended to random dot patterns in the 100% rotating regular motion and then noted positional changes of these patterns. Therefore, it is possible that the rotating stimuli may have activated attention based position tracking. To determine whether any position tracking is necessary for motion direction and static tone association in a more controlled way, we used reverse-phi motion in the association phase of Experiment 2.

As shown in **Figures 4 and 5**, the audiovisual association had different effects on the subsequent perceived directions of regular and reverse-phi motions [three-way repeated-measures ANOVA on PSE shifts motion type, test phase and sound condition

as factors, three-way interaction $F(1,5) = 8.601$, $p < 0.05$]. However, the main effects of motion type [$F(1,5) = 0.122$, $p = 0.741$], test phase [$F(1,5) = 0.152$, $p = 0.713$] and sound condition [$F(1,5) = 4.491$, $p = 0.088$] were not significant. After the association phase, static tones biased the perceived direction of the reverse-phi motion in favor of the paired presentations of perceived directions of reverse-phi motion and static tones (**Figure 4D**). As opposed to the pre-association test phase [$F(1,5) = 1.824$, $p = 0.235$], this led to a significant difference between two sound conditions in the post-association [**Figure 5B**, simple main effect of sound $F(1,5) = 8.954$, $p < 0.05$]. On the other hand, static tones did not cause any consistent change in the perceived direction of regular motion [simple main effect of sound: pre-association test $F(1,5) = 1.794$, $p = 0.238$; post-association test $F(1,5) = 0.336$, $p = 0.587$]. As in Experiment 1, we did not find any significant difference between the bright and dark conditions of regular motion (Supplementary Table S2).



Motion Direction Discrimination

The paired presentation of static tones with visual motion directed toward a given direction during association phase may also change subsequent direction discrimination thresholds. More specifically, in the post association test phase, static tones may increase the direction discrimination performance and decrease motion discrimination thresholds (i.e., steeper psychometric curves) for each sound condition relative to visual-only. We tested this possibility by using the changes in slope values (SDs) shown in **Figure 6**. A three-way repeated measures ANOVA (motion type, test phase and sound condition as factors) revealed no significant factor and interaction (Supplementary Table S3) for Experiment 1 (**Figures 6A,C**). For the second experiment, we found only a significant three-way interaction [Supplementary Table S4, $F(1,5) = 40.102$, $p < 0.01$]. Follow-up tests revealed only a simple main effect of sound [$F(1,5) = 7.309$, $p < 0.05$] in the post-association test of reverse-phi motion (**Figures 4D** and **6D**). This suggests that the leftward shift of the psychometric function for the CCW sound condition (**Figure 4D**) is possibly confounded by the significant increase in the slope value. Overall, our results do not provide any consistent

sound specific decrease in slope values and increase in direction discrimination performance due to audiovisual associations. If anything, the slope values tend to increase and direction discrimination performance tends to decrease after the association phase (for additional tests on raw slope values see Supplementary Tables S5 and S6).

Discussion

Associations have been considered to play a role to disambiguate motion stimuli in noisy environments (Albright, 2012). Several studies have reported a compelling influence of audiovisual associations on ambiguous visual motion (Teramoto et al., 2010; Hidaka et al., 2011). After a brief association phase, static sounds can bias the perception of ambiguous visual motion and even induce a motion perception from static flickering objects. The mechanisms underlying these association induced effects have not been understood within the framework of motion processing hierarchy. In this study, we investigated whether audiovisual associations and their subsequent influences on motion perception were dependent on the involvement of higher-order attention-dependent position tracking mechanisms. We used regular and also reverse-phi motions to isolate low-level motion energy and to minimize the involvement of any explicit attentive tracking. After paired presentations of regular motion directions and static tones for around 8 min, static tones significantly biased the perceived directions of both regular and reverse-phi motions in favor of the exposed audiovisual pairing. In a follow-up experiment, we used reverse-phi motion during the association phase. We found that the association established between the (perceived) directions of reverse-phi motion and static tones affected only the subsequent perception of reverse-phi motion. We reasoned that if positional tracking were necessary for the association to be established and to impact the motion perception, then the perception of these low-level motion types should have been unaffected by static tones in the post-association test sessions as observed in the pre-association data. However, our results indicate significant changes in low-level motion perception (especially reverse-phi motion attributed to low-level motion energy mechanisms). It is also unlikely that the observed effects of audiovisual associations would simply be explained by response/decisional biases. If any decisional bias led to the observed changes in behavioral results, the association phase should have affected both motion types in a similar way. This is not the case when the reverse-phi motion is used for the association phase. Accordingly, the findings presented here suggest that audiovisual associations are not restricted to high-level attention based motion system, and that early-level motion processing has some potential role. In what follows, we discuss our findings with respect to their implications to visual motion processing, audiovisual (more generally multisensory) associations and interactions.

Regular Versus Reverse-phi Motion Types

When we compare the outcomes from both experiments, it is obvious that the association phase with regular motion

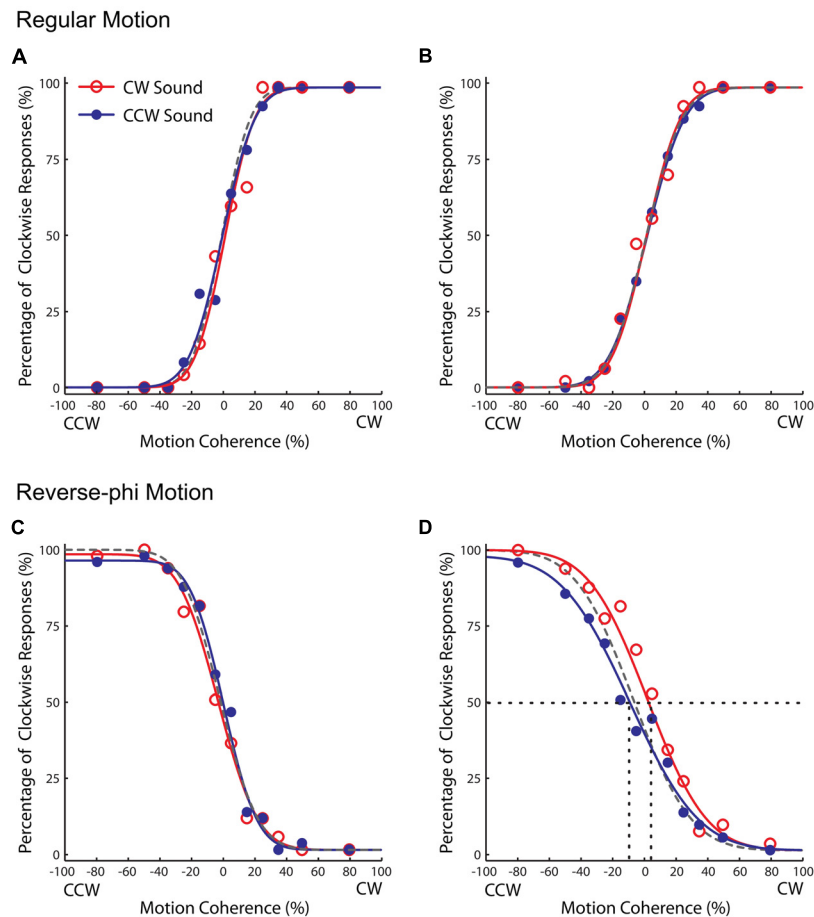


FIGURE 4 | Results of Experiment 2 (reverse-phi motion used in the association phase). Group averaged data ($n = 6$, five were naïve observers) for different experimental conditions. Top panel: pre-association (A) and post-association (B) data for regular motion. Bottom panel:

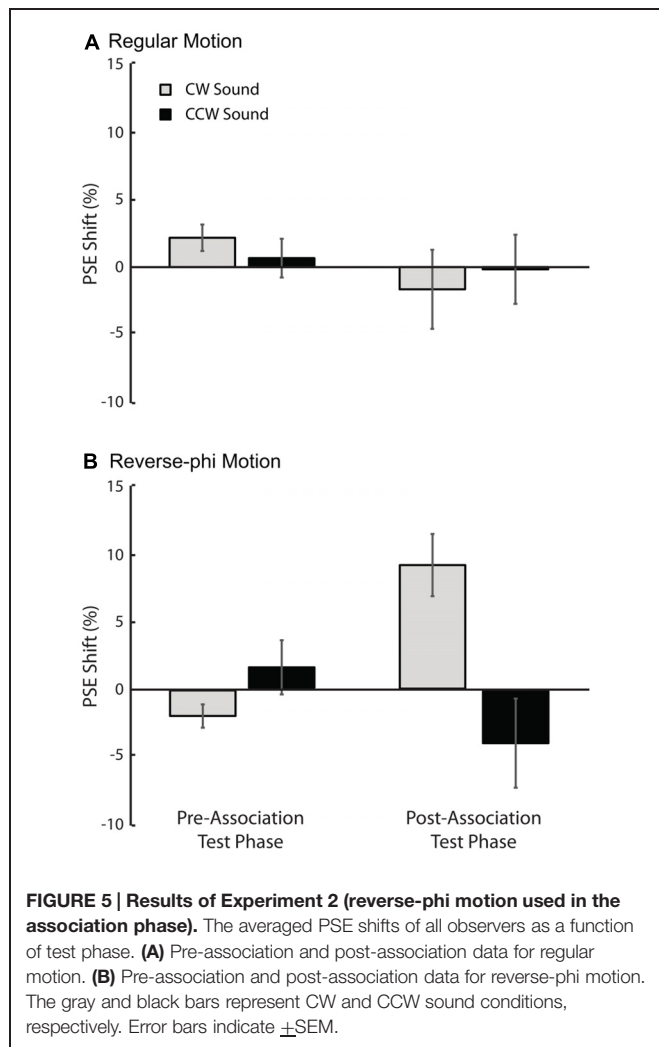
pre-association (C) and post-association (D) data for reverse-phi motion. Other conventions are the same as those in Figure 2. Each plot indicates the proportion of trials in which the visual stimulus was judged to rotate in clockwise direction.

(Experiment 1: Figure 3) led to greater perceptual biases than the one with reverse-phi motion (Experiment 2: Figure 5). As mentioned above, the association phase with reverse-phi motion did not even bias the perceived direction of regular motion. Previous studies showed that adaptation to different types of visual motion can result in aftereffects with distinct characteristics. These distinctions have been mostly explained by differences between the mechanisms involved in processing visual motion types (Mather et al., 1998). Even though the tracking of individual dot positions was difficult for both regular and reverse-phi motion types in our experiments, it is still possible that regular phi motion might engage attention based position tracking more than the reverse-phi motion does. This minimal involvement of high-level attention based system may result in a more effective association phase for the regular phi motion. Another possibility is that there might be differences in the way these motion types drive the low-level motion detectors. More recently, low-level detection of regular and reverse-phi motions have been found to be different and it is suggested that these motion types drive low-level motion detectors in different ways

(Bours et al., 2007). The differences found between low-level motion detection of regular and reverse-phi motion could also account for the differences observed in the audiovisual formed associations based on these two different motion types used in our study. Future studies are required to characterize the differences between the two audiovisual association types and to understand how they are linked to the motion processes involved.

Low-Level and High-Level Processes in Visual Motion

Our findings demonstrate for the first time that the association established between visual motion directions and static tones impact the perceived direction of visual motion even when attentive tracking is ruled out. An interesting issue is whether these behavioral findings provide some information about which motion areas are involved in these audiovisual associations. Using functional imaging and motion stimuli which selectively tap into different motion systems, several studies identified distinct cortical areas involved in low-level pre-attentive and high-level



attention based motion processing (Claeys et al., 2003; Ho and Giaschi, 2009a). Although lower visual areas (V1, V3A, and hMT+) appeared to be sensitive to all types of visual motion, several cortical areas were selectively activated by the motion types engaging low-level pre-attentive and high-level attentive tracking mechanisms preferentially (Claeys et al., 2003). Building from these findings, Claeys et al. (2003) identified the neuroanatomical substrates for these motion-processing systems in the human brain: the contralateral low-level motion energy system, extending from hMT+ into dorsal inferior parietal sulcus (IPS) and superior temporal sulcus (STS), and the bilateral higher level saliency-based system in the inferior parietal lobe (IPL). Moreover, IPS activation for the visual motion that increase the involvement of attentive tracking of salient visual features has been reported as significantly higher than for the one restricting any explicit attentive tracking (low-level visual motion), suggesting that this region represents a further stage than V3A and hMT+ in the motion processing hierarchy (Ho and Giaschi, 2009a). Based on these studies, our psychophysical results suggest that the influences of audiovisual associations on motion processing are not restricted to higher-order motion areas (e.g., IPL) mediating

attention-based position tracking. They imply that lower visual motion areas (V1, V3A, and hMT+) are probably involved in audiovisual associations.

Our findings are in line with converging evidence showing that associative plasticity is not restricted to high-level visual motion processing and also present for neurons at early processing stages (Albright, 2012). By using a visual paired association task to link the direction of a static arrow with the direction of a random dot displays, Schlack and Albright (2007) found that many MT neurons exhibit selectivity for the direction of the static arrow. This property was not seen before the association between the static arrow and motion direction was formed. In addition to the relatively early-level plasticity in MT neurons, Fitzgerald et al. (2011) reported more profound changes in the selectivity of LIP neurons (Lateral Intraparietal cortex, part of monkey IPS) not only by associations between motion directions but also by other learned associations between pairs of arbitrarily chosen static shapes. More importantly, in a recent study, Zilber et al. (2014) trained human observers for 80 min to discriminate visual motion directions. They observed a great improvement in performance when the visual stimuli were presented with coherent acoustic textures, but not when presented with auditory noise. The MEG data collected in tandem revealed cortical plasticity in hMT+ reflecting the perceptual improvement associated with coherent auditory stimuli.

There have been numerous studies examining interactions between sound and visual motion (for a review see Soto-Faraco et al., 2003). Several illusions have demonstrated the strong influence of auditory stimuli on the perception of visual motion (e.g., Sekuler et al., 1997; Freeman and Driver, 2008; Hidaka et al., 2009). Of more direct relevance to our current study, many recent studies have suggested that the temporal and spatial information provided by audition is used at early stages of motion processing (Alink et al., 2008; Scheef et al., 2009; Kafaligonul and Stoner, 2010, 2012; Ogawa and Macaluso, 2013; von Saldern and Noppeney, 2013). That is to say, auditory information may interact with spatiotemporal processing of motion (low-level motion energy) at lower visual areas which have been traditionally thought to be unisensory (V1 and/or MT). The behavioral results presented here extend this recent view by showing that not only explicit stimulus properties of sound but also the audiovisual associations acquired through a brief exposure phase are also involved in shaping low-level visual motion perception, and they may be used at early stages of motion processing.

Potential Neural Mechanisms

Our results do not preclude high-level processes (e.g., attention-dependent positional tracking), but support the view that low-level motion processing has a role in multisensory associations. The involvement of lower motion areas can take place in various forms. In general, associations are believed to be mediated by a strengthening of connectivity between the relevant neural representations (Albright, 2012). Neuroanatomical studies show some evidence of sparse connections between auditory and primary visual cortices (Falchier et al., 2002; Clavagnier et al., 2004; Cappe and Barone, 2005). It has been suggested that these connections

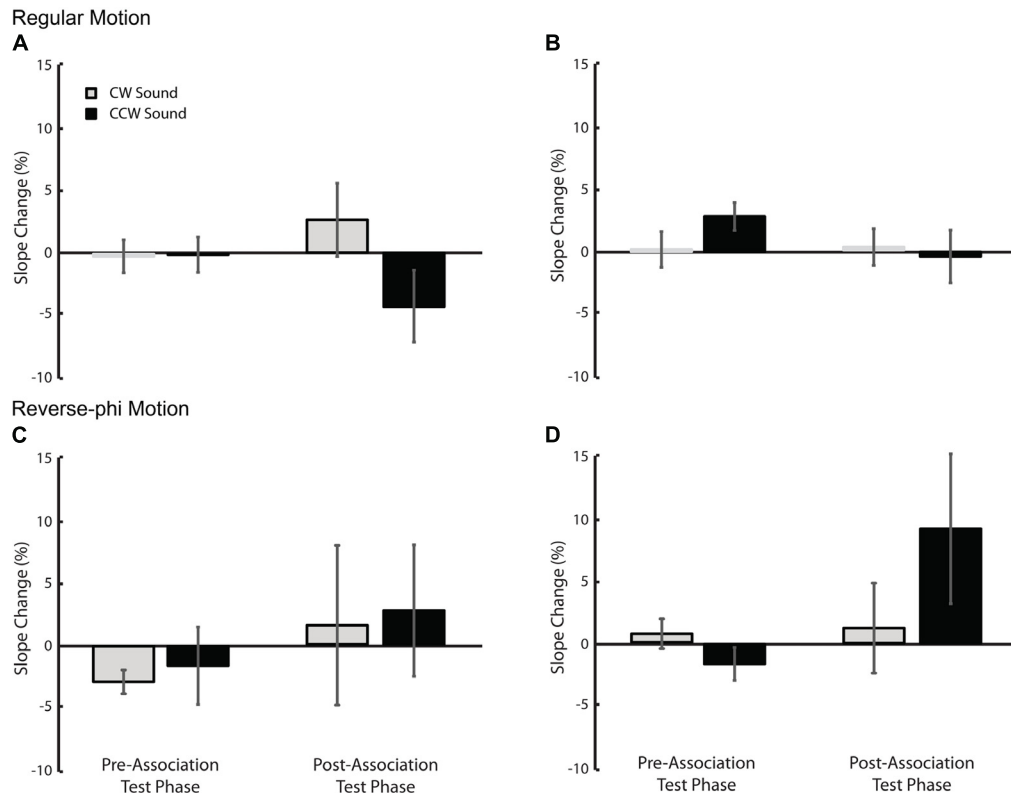


FIGURE 6 | The changes in slope values (in units of motion coherence level) relative to the visual-only condition. Top panel: Experiment 1 (A) and Experiment 2 (B) regular motion data. Bottom panel: Experiment 1 (C) and Experiment 2 (D) reverse-phi motion data. Error bars indicate \pm SEM. Other conventions are the same as those in Figures 3 and 5.

may mediate early-level crossmodal interactions and auditory- or visual-driven phase resetting found in primary sensory cortices (Schroeder and Foxe, 2005; Kayser et al., 2008; Mercier et al., 2013). Accordingly, repeated exposure to static tones and moving stimuli in different directions may result in a strengthening of this early “crosstalk” between auditory and visual cortical areas (i.e., the lateral feedforward connections between auditory and lower visual motion areas). Vetter et al. (2014) have recently found that sounds can also induce mental imagery and hence, activate early visual cortex by feedback from non-retinal resources. In a similar way, static sounds may eventually elicit a pictorial recall of motion direction due to the paired presentations of static sounds with motion directions many times (the brief audiovisual association phase). Therefore, the influence of audiovisual associations on low-level visual motion may be mediated by a strengthening of top-down feedback connections from multisensory association and/or non-sensory areas to lower visual areas (Petro et al., 2014).

Conclusion

In conclusion, the present study demonstrates that the association established between visual motion directions and static tones significantly affects low-level visual motion perception.

These findings, in conjunction with a variety of related converging findings, suggest that audiovisual associations are not restricted to later stages of motion processing and lower visual motion areas has some potential role. Future neuroimaging and neurophysiological studies are needed to explicitly test whether audiovisual associations are restricted to higher-order visual areas and whether the audiovisual associations are mediated by feed-forward lateral connections, top-down modulatory feedback or both.

Acknowledgments

We thank Michelle Adams, Huseyin Boyaci, and Jedediah Allen for discussions and comments on the manuscript. This work was supported by the Scientific and Technological Research Council of Turkey (TUBITAK Grants 112C010 and 113K547). Some portion of the results were presented at 37th European Conference on Visual Perception.

Supplementary Material

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fnint.2015.00026/abstract>

References

- Adelson, E. H., and Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *J. Opt. Soc. Am. A* 2, 284–299. doi: 10.1364/JOSAA.2.000284
- Albright, T. D. (2012). On the perception of probable things: neural substrates of associative memory, imagery, and perception. *Neuron* 74, 227–245. doi: 10.1016/j.neuron.2012.04.001
- Alink, A., Singer, W., and Muckli, L. (2008). Capture of auditory motion by vision is represented by an activation shift from auditory to visual motion cortex. *J. Neurosci.* 28, 2690–2697. doi: 10.1523/JNEUROSCI.2980-07.2008
- Anstis, S. M. (1970). Phi movement as a subtraction process. *Vision Res.* 10, 1411–1430. doi: 10.1016/0042-6989(70)90092-1
- Anstis, S. M. (1980). The perception of apparent movement. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 290, 153–168. doi: 10.1098/rstb.1980.0088
- Azzopardi, P., and Hock, H. S. (2011). Illusory motion perception in blindsight. *Proc. Natl. Acad. Sci. U.S.A.* 108, 876–881. doi: 10.1073/pnas.1005974108
- Bours, R. J. E., Kroes, M. C. W., and Lankheet, M. J. M. (2007). The parallel between reverse-phi and motion aftereffects. *J. Vis.* 7, 1–10. doi: 10.1167/7.11.8
- Braddick, O. J. (1980). Low-level and high level processes in apparent motion. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 290, 137–151. doi: 10.1098/rstb.1980.0087
- Brainard, D. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436. doi: 10.1163/156856897X00357
- Britten, K. H., Shadlen, M. N., Newsome, W. T., and Movshon, J. A. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J. Neurosci.* 12, 4745–4765.
- Cappe, C., and Barone, P. (2005). Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. *Eur. J. Neurosci.* 22, 2886–2902. doi: 10.1111/j.1460-9568.2005.04462.x
- Cavanagh, P. (1992). Attention-based motion perception. *Science* 257, 1563–1565. doi: 10.1126/science.1523411
- Claydes, K. G., Lindsey, D. T., De Schutter, E., and Orban, G. A. (2003). A higher order motion region in human inferior parietal lobule: evidence from fMRI. *Neuron* 40, 451–452. doi: 10.1016/S0896-6273(03)00590-7
- Clavagnier, S., Falchier, A., and Kennedy, H. (2004). Long distance feedback projections to area V1: implications for multisensory integration, spatial awareness, and visual consciousness. *Cogn. Affect. Behav. Neurosci.* 4, 117–126. doi: 10.3758/CABN.4.2.117
- Culham, J. C., Verstraten, F. A. J., Ashida, H., and Cavanagh, P. (2000). Independent aftereffects of attention and motion. *Neuron* 28, 607–615. doi: 10.1016/S0896-6273(00)00137-9
- Emerson, R. C., Citron, M. C., Vaughn, W. J., and Klein, S. A. (1987). Nonlinear directionally selective subunits in complex cells of cat striate cortex. *J. Neurophysiol.* 58, 33–65.
- Falchier, A., Clavagnier, S., Barone, P., and Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. *J. Neurosci.* 22, 5749–5759.
- Fitzgerald, J. K., Freedman, D. J., and Assad, J. A. (2011). Generalized associative representations in parietal cortex. *Nat. Neurosci.* 14, 1075–1079. doi: 10.1038/nn.2878
- Freeman, E., and Driver, J. (2008). Direction of visual apparent motion driven solely by timing of a static sound. *Curr. Biol.* 18, 1262–1266. doi: 10.1016/j.cub.2008.07.066
- Haijiang, Q., Saunders, J. A., Stone, R. W., and Backus, B. T. (2006). Demonstration of cue recruitment: change in visual appearance by means of Pavlovian conditioning. *Proc. Natl. Acad. Sci. U.S.A.* 98, 483–488. doi: 10.1073/pnas.0506728103
- Hidaka, S., Manaka, Y., Teramoto, W., Sugita, Y., Miyauchi, R., Gyoba, J., et al. (2009). Alternation of sound location induces visual motion perception of a static object. *PLoS ONE* 4:e8188. doi: 10.1371/journal.pone.0008188
- Hidaka, S., Teramoto, W., Kobayashi, M., and Sugita, Y. (2011). Sound-contingent visual motion aftereffect. *BMC Neurosci.* 12:44. doi: 10.1186/1471-2202-12-44
- Ho, C. S., and Giaschi, D. E. (2009a). Low- and high-level first-order random-dot kinematograms: evidence from fMRI. *Vision Res.* 49, 1814–1824. doi: 10.1016/j.visres.2009.04.018
- Ho, C., and Giaschi, D. (2009b). Low- and high-level motion perception deficits in anisometropic and strabismic amblyopia: evidence from fMRI. *Vision Res.* 49, 2891–2901. doi: 10.1016/j.visres.2009.07.012
- Hock, H. S., and Nichols, D. F. (2013). The perception of object vs. object-less motion. *Atten. Percept. Psychophys.* 75, 726–737. doi: 10.3758/s13414-013-0441-1
- Kafaligonul, H., and Stoner, G. R. (2010). Auditory modulation of visual apparent motion with short spatial and temporal intervals. *J. Vis.* 10, 1–13. doi: 10.1167/10.12.31
- Kafaligonul, H., and Stoner, G. R. (2012). Static sound timing alters sensitivity to low-level visual motion. *J. Vis.* 12, 1–9. doi: 10.1167/12.11.2
- Kayser, C., Petkov, C. I., and Logothetis, N. K. (2008). Visual modulation of neurons in auditory cortex. *Cereb. Cortex* 18, 1560–1574. doi: 10.1093/cercor/bhm187
- Kim, R., Seitz, A., Feenstra, H., and Shams, L. (2009). Testing assumptions of statistical learning: is it implicit and long-term? *Neurosci. Lett.* 461, 145–149. doi: 10.1016/j.neulet.2009.06.030
- Kobayashi, M., Teramoto, W., Hidaka, S., and Sugita, Y. (2012a). Indiscernible sounds determine the direction of visual motion. *Sci. Rep.* 2:365. doi: 10.1038/srep00365
- Kobayashi, M., Teramoto, W., Hidaka, S., and Sugita, Y. (2012b). Sound frequency and aural selectivity in sound-contingent visual motion aftereffect. *PLoS ONE* 7:e36803. doi: 10.1371/journal.pone.0036803
- Krekelberg, B., and Albright, T. D. (2005). Motion mechanisms in macaque MT. *J. Neurophysiol.* 93, 2908–2921. doi: 10.1152/jn.00473.2004
- Lu, Z. L., and Sperling, G. (1995a). Attention-generated apparent motion. *Nature* 377, 237–239. doi: 10.1038/377237a0
- Lu, Z. L., and Sperling, G. (1995b). The functional architecture of human visual motion perception. *Vision Res.* 35, 2697–2722. doi: 10.1016/0042-6989(95)00025-U
- Mather, G., Verstraten, F., and Anstis, S. (1998). *The Motion After-Effect: A Modern Perspective*. Cambridge, MA: MIT Press.
- Mercier, M. R., Foxe, J. J., Fiebelkorn, I. C., Butler, J. S., Schwartz, T. H., and Molholm, S. (2013). Auditory-driven phase reset in visual cortex: human electrocorticography reveals mechanisms of early multisensory integration. *Neuroimage* 79, 19–29. doi: 10.1016/j.neuroimage.2013.04.060
- Murphy, A. P., Leopold, D. A., and Welchman, A. E. (2014). Perceptual memory drives learning of retinotopic biases for bistable stimuli. *Front. Psychol.* 5:60. doi: 10.3389/fpsyg.2014.00060
- Ogawa, A., and Macaluso, E. (2013). Audio-visual interactions for motion perception in depth modulate activity in visual area V3A. *Neuroimage* 71, 158–167. doi: 10.1016/j.neuroimage.2013.01.012
- Pelli, D. (1997). The videotoolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* 10, 437–442. doi: 10.1163/156856897X00366
- Petro, L. S., Vizioli, L., and Muckli, L. (2014). Contributions of cortical feedback to sensory processing in primary visual cortex. *Front. Psychol.* 5:1223. doi: 10.3389/fpsyg.2014.01223
- Pilly, P. K., and Seitz, A. R. (2009). What a difference a parameter makes: a psychophysical comparison of random dot motion algorithms. *Vision Res.* 49, 1599–1612. doi: 10.1016/j.visres.2009.03.019
- Sato, T. (1998). “Dmax: relations to low- and high-level motion processes,” in *High-Level Motion Processing, Computational, Neurobiological, and Psychophysical Perspectives*, ed. T. Watanabe (Boston: MIT Press), 115–151.
- Scheef, L., Boecker, H., Daamen, M., Fehse, U., Landsberg, M. W., Granath, D. O., et al. (2009). Multimodal motion processing in area V5/MT: evidence from an artificial class of audio visual events. *Brain Res.* 1252, 94–104. doi: 10.1016/j.brainres.2008.10.067
- Schlack, A., and Albright, T. D. (2007). Remembering visual motion: neural correlates of associative plasticity and motion recall in cortical area MT. *Neuron* 53, 881–890. doi: 10.1016/j.neuron.2007.02.028
- Schlack, A., Vivian, V., and Albright, T. D. (2007). Altering motion perception by motion – colour pairing. *Perception* 36, 1–235.
- Schroeder, C. E., and Foxe, J. (2005). Multisensory contributions to low-level, ‘unisensory’ processing. *Curr. Opin. Neurobiol.* 15, 454–458. doi: 10.1016/j.conb.2005.06.008
- Schutz, A. C., Braun, D. I., Movshon, J. A., and Gegenfurtner, K. R. (2010). Does the noise matter? Effects of different kinematogram types on smooth pursuit eye movements and perception. *J. Vis.* 10:26. doi: 10.1167/10.13.26

- Seitz, A. R., Kim, R., van Wassenhove, V., and Shams, L. (2007). Simultaneous and independent acquisition of multisensory and unisensory associations. *Perception* 36, 1445–1453. doi: 10.1068/p5843
- Sekuler, R., Sekuler, A. B., and Lau, R. (1997). Sound alters visual motion perception. *Nature* 385:308. doi: 10.1038/385308a0
- Shams, L., Wozny, D. R., Kim, R., and Seitz, A. (2011). Influences of multisensory experience on subsequent unisensory processing. *Front. Psychol.* 2:264. doi: 10.3389/fpsyg.2011.00264
- Soto-Faraco, S., Kingstone, A., and Spence, C. (2003). Multisensory contributions to the perception of motion. *Neuropsychologia* 41, 1847–1862. doi: 10.1016/S0028-3932(03)00185-4
- Teramoto, W., Hidaka, S., and Sugita, Y. (2010). Sounds move a static visual object. *PLoS ONE* 5:e12255. doi: 10.1371/journal.pone.0012255
- Vetter, P., Smith, F. W., and Muckli, L. (2014). Decoding sound and imagery content in early visual cortex. *Curr. Biol.* 24, 1256–1262. doi: 10.1016/j.cub.2014.04.020
- von Saldern, S., and Noppeney, U. (2013). Sensory and striatal areas integrate auditory and visual signals into behavioral benefits during motion discrimination. *J. Neurosci.* 33, 8841–8849. doi: 10.1523/JNEUROSCI.3020-12.2013
- Wichmann, F. A., and Hill, N. J. (2001a). The psychometric function: I. Fitting, sampling and goodness-of-fit. *Percept. Psychophys.* 63, 1293–1313. doi: 10.3758/BF03194544
- Wichmann, F. A., and Hill, N. J. (2001b). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Percept. Psychophys.* 63, 1314–1329. doi: 10.3758/BF03194545
- Zilber, N., Ciuciu, P., Gramfort, A., Azizi, L., and van Wassenhove, V. (2014). Supramodal processing optimizes visual perceptual learning and plasticity. *Neuroimage* 93, 32–46. doi: 10.1016/j.neuroimage.2014.02.017

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Kafaligonul and Oluk. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.